EasyChair Preprint

№ 15562

Using ChatGPT to Build Healthcare Agents -an Empirical Study

Hassan Juma, Ibrahim Mohammed and Sarah Zhao

December 11, 2024

Using ChatGPT to Build Healthcare Agents -An Empirical Study

Hassan Juma[1], Ibrahim Mohammed[2], Sarah Zhao[3]

1. University of Dodoma (UDOM)

2. Muhimbili University of Health and Allied Sciences (MUHAS)

3. Swiss Hotel Management School (SHMS)

Introduction

The integration of ChatGPT into healthcare agent development represents a significant advancement in medical technology and patient care delivery(Abbasian, M)(Li, X)(Ren, Z). As healthcare systems increasingly embrace artificial intelligence, the development of intelligent healthcare agents powered by large language models has emerged as a promising frontier for improving medical services and patient outcomes. These AI-driven healthcare agents can assist medical professionals in various tasks, from patient communication to clinical decision support, while maintaining strict privacy and security standards(Abbasian, M). The implementation of ChatGPT in healthcare contexts requires careful consideration of ethical implications, data protection, and regulatory compliance(Li, X). Recent developments in federated learning approaches, such as FedAPT, have demonstrated the potential for collaborative learning while preserving patient privacy across different healthcare domains(Abbasian, M).

Technical Components of Healthcare Agent Development

Data Augmentation and Model Training

The integration of ChatGPT in healthcare agent development has revolutionized the approach to medical dataset enhancement and model training processes(Abbasian, M). By leveraging ChatGPT's advanced language understanding capabilities, healthcare organizations can significantly augment their existing medical datasets with synthetically generated, yet clinically relevant scenarios. This process involves careful consideration of the healthcare domain's intricacies, ensuring that the generated data maintains high clinical accuracy while expanding the training corpus. The augmented datasets enable more robust training of healthcare agents, particularly in understanding complex medical queries and providing personalized responses.

Integration Framework

The development of an effective integration framework for ChatGPT-based healthcare agents requires a sophisticated architectural approach that ensures seamless interaction between various system components(Abbasian, M). The framework must support multiple functionalities including query understanding, knowledge possession, health data access, and analytical capabilities. The architecture incorporates specialized

modules for processing multimodal inputs, managing personalized health information, and accessing current medical literature. This integration framework emphasizes the importance of maintaining continuous updates with advancing healthcare technology while ensuring robust security protocols for handling sensitive medical data.

Empirical Performance Analysis

In a comprehensive evaluation of ChatGPT's capabilities in healthcare applications, empirical evidence demonstrates varying levels of effectiveness across different medical tasks and specialties. According to recent performance metrics, ChatGPT and its variants show promising results in medical knowledge application, though with notable variations across different tools and approaches(Schmidgall, S). The data indicates that Claude 3.5 achieved the highest average performance at 53.2% across medical specialties, followed by GPT-4 at 41.5%. When examining specific agent tools, the Notebook approach yielded the best results for Claude 3.5 at 56.1%, while Adaptive RAG with web sources showed consistent performance at 52.4%(Schmidgall, S). In specialized medical tasks such as named entity recognition and relation extraction, the models demonstrated varying capabilities, with performance metrics ranging from 30.7% to 67.5% in different applications(Xie, Q).

Ethical Considerations and Challenges

Privacy and Security Concerns

In the rapidly evolving landscape of healthcare AI applications, privacy and security concerns stand at the forefront of ethical considerations(Haltaufderheide, J). The implementation of ChatGPT-based healthcare agents necessitates robust data protection measures to safeguard sensitive patient information. Healthcare organizations must establish comprehensive security protocols that align with regulatory requirements while ensuring patient confidentiality remains uncompromised. The challenge lies in balancing the need for data accessibility to train and improve AI systems with the imperative to protect individual privacy rights.

Bias and Accuracy Issues

The development of AI-driven healthcare agents faces significant challenges related to bias and accuracy in medical responses(Haltaufderheide, J). Large Language Models (LLMs) may inadvertently perpetuate existing biases present in their training data, potentially leading to disparities in healthcare recommendations across different demographic groups. Also, the LLMs may generate bias caused by the user input, namely prompt bias(Xu, Z). The accuracy of AI-generated medical information remains a critical concern, as these systems can sometimes produce convincingly presented but inaccurate content(Haltaufderheide, J). This challenge is particularly significant given the high-stakes nature of healthcare decisions. The need for human oversight becomes paramount, as highlighted by recurring ethical concerns regarding fairness, transparency, and non-maleficence(Haltaufderheide, J).

Future Directions and Recommendations

Developing healthcare agents using ChatGPT presents several promising avenues for future advancement and improvement. A critical area for enhancement lies in the sophisticated alignment of different modalities within healthcare systems(Niu, Q). Future research should focus on developing novel architectures and training methodologies that can better capture the intricate relationships between various medical data types, leading to more accurate and robust clinical predictions. The interpretability and explainability of Multi-modal Large Language Models (MLLMs) remain paramount concerns that need addressing, particularly in clinical settings where transparency is crucial for building trust and ensuring adoption(Niu, Q).

To prevent potential power imbalances and exploitation(Zhou, J), it is essential to distribute healthcare information and communication systems across multiple platforms, promoting competition and diversity in healthcare technology. This approach ensures that no single entity maintains excessive control over healthcare information flow while preserving individual access to critical medical information. The ethical considerations in healthcare agent development demand particular attention(Zhou, J), especially concerning the moral behavior of AI agents in healthcare settings.

Human oversight plays a vital role in providing context and ethical judgment(Ferrara, E), helping to identify and address potential biases, errors, or unintended consequences in healthcare applications. The integration of these systems into clinical practice has shown promising results in enhancing patient outcomes through data-driven decision-making(Nazi, Z). Moving forward, continued research into bias identification and mitigation methods will be crucial for advancing equitable and inclusive AI healthcare solutions. The focus should remain on developing systems that prioritize patient care while maintaining the highest standards of ethical practice and professional responsibility.

References

[1] Abbasian, M., Azimi, I., Rahmani, A. M., & Jain, R. (2023). Conversational Health Agents (CHAs) are interactive systems designed to enhance personal healthcare services by engaging in empathetic conversations and processing multimodal data. https://arxiv.org/pdf/2310.02374v2

[2] Li, X., Peng, L., Wang, Y., & Zhang, W. (2024). This survey explores the transformative impact of foundation models (FMs) in artificial intelligence, focusing on their integration with federated learning (FL) for advancing biomedical research. https://arxiv.org/pdf/2405.06784

[3] Ren, Z., Zhan, Y., Yu, B., Ding, L., & Tao, D. (2024). Healthcare copilot: Eliciting the power of general llms for medical consultation. *arXiv preprint arXiv:2402.13408*. https://arxiv.org/pdf/2402.13408

[4] Abbasian, M., Azimi, I., Rahmani, A. M., & Jain, R. (2023). Conversational Health Agents (CHAs) are interactive systems that provide healthcare services, such as assistance, self-awareness, and diagnosis. https://arxiv.org/html/2310.02374v3

[5] Abbasian, M., Azimi, I., Rahmani, A. M., & Jain, R. (2024). Conversational Health Agents (CHAs) are interactive systems that provide healthcare services, such as assistance and diagnosis. https://arxiv.org/html/2310.02374v5

[6] Schmidgall, S., Ziaei, R., Harris, C., Reis, E., Jopling, J., & Moor, M. (2024). Evaluating large language models (LLM) in clinical scenarios is crucial to assessing their potential clinical utility. https://arxiv.org/abs/2405.07960

[7] Xie, Q., Chen, Q., Chen, A., Peng, C., Hu, Y., Lin, F., Peng, X., Huang, J., Zhang, J., Keloth, V., Zhou, X., Qian, L., He, H., Shung, D., Ohno-Machado, L., Wu, Y., Xu, H., Bian, J. (2024). Recent advancements in large language models (LLMs) like ChatGPT and LLaMA show promise in medical applications, yet challenges remain in medical language comprehension. https://arxiv.org/pdf/2402.12749

[8] Haltaufderheide, J., & Ranisch, R. (2024). With the introduction of ChatGPT, Large Language Models (LLMs) have received enormous attention in healthcare. https://arxiv.org/html/2403.14473v1

[9] Haltaufderheide, J., & Ranisch, R. (2024). With the introduction of ChatGPT, Large Language Models (LLMs) have received enormous attention in healthcare. https://arxiv.org/pdf/2403.14473

[10] Xu, Z., Peng, K., Ding, L., Tao, D., & Lu, X. (2024). Take Care of Your Prompt Bias! Investigating and Mitigating Prompt Bias in Factual Knowledge Extraction. https://aclanthology.org/2024.lrec-main.1352.pdf

[11] Niu, Q., Chen, K., Li, M., Feng, P., Bi, Z., Yan, L. K., Zhang, Y., Yin, C. H., Fei, C., Liu, J., Peng, B., Wang, T., Wang, Y., Chen, S., Liu, M. (2024). Large Language Models (LLMs) have rapidly evolved from text-based systems to multimodal platforms, significantly impacting various sectors including healthcare. https://arxiv.org/html/2410.01812v5

[12] Zhou, J., Müller, H., Holzinger, A., & Chen, F. (2023). Large language models, e.g. ChatGPT are currently contributing enormously to make artificial intelligence even more popular, especially among the general population. https://arxiv.org/abs/2305.10646

[13] Zhou, J., Müller, H., Holzinger, A., & Chen, F. (2023). Large language models, e.g. ChatGPT are currently contributing enormously to make artificial intelligence even more popular, especially among the general population. https://arxiv.org/pdf/2305.10646

[14] Ferrara, E. (2023). As the capabilities of generative language models continue to advance, the implications of biases ingrained within these models have garnered increasing attention from researchers, practitioners, and the broader public. https://arxiv.org/abs/2304.03738

[15] Nazi, Z. A., & Peng, W. (2024). The deployment of large language models (LLMs) within the healthcare sector has sparked both enthusiasm and apprehension. https://arxiv.org/html/2401.06775v2